ELSEVIER Discrete Mathematics 148 (1996) 311-316

DISCRETE MATHEMATICS

Note

On the evolution of a random tournament ☆

Tomasz Łuczak [a,*,1], Andrzej Ruciński [b], Jacek Gruszka [c]

[a] Mathematical Institute of the Polish Academy of Sciences, Poznań, Poland
[b] Department of Discrete Mathematics, Adam Mickiewicz University, Poznań, Poland
[c] Institute of Mathematics, Technical University, Poznań, Poland

Received 5 May 1993; revised 26 May 1994

1. Introduction

Let $T(n,p)$ be a tournament with the vertex set $[n] = \{1,2,\dots,n\}$ such that the pair $\{i,j\}$, $i<j$, is oriented from i to j with probability p, independently for each $i,j \in [n]$. The particular case of $p=\frac{1}{2}$ was investigated by several authors (for further references see [5]). For arbitrary p, according to our knowledge, this model was introduced in [1]. In this note we study asymptotic properties of $T(n,p)$, as $n \to \infty$ and parameter p may vary as a function of n. We determine the threshold functions for small subgraphs of $T(n,p)$, i.e. for those whose size does not depend on n. Then we characterize component structure of $T(n,p)$ and deduce from it the length of the longest cycle in $T(n,p)$. As a corollary we get the threshold function for the existence of a Hamilton cycle in $T(n,p)$.

Throughout the note paths and cycles are always meant to be directed and the connectivity should be understood as the strong connectivity of digraphs.

2. Small subgraphs

A subgraph of a tournament is necessarily an antisymmetric digraph, i.e. it does not contain arcs (i,j) and (j,i) at the same time. On the other hand, if $p \neq 0,1$ then each antisymmetric digraph of size at most n appears in $T(n,p)$ with positive

☆ Authors were partially supported by KBN grant 2 1087 91 01.
* Corresponding author.
[1] On leave from Adam Mickiewicz University, Poznań, Poland

SSDI 0012-365X(94)00279-7

probability. In this section we investigate the limit $\lim_{n\to\infty} \mathrm{Prob}(T(n,p)\supset D)$, where D is an arbitrary but fixed antisymmetric digraph. From now on by a digraph we shall always mean an antisymmetric one.

Consider first the case of an acyclic digraph D. Since every acyclic digraph can be trivially extended to the transitive tournament on the same vertex set, and the transitive tournament of order k is contained in *every* tournament on $n \geqslant R(k,k)$ vertices, where $R(k,k)$ is the Ramsey number (see [2, p. 18]), so, in this case, $\mathrm{Prob}(T(n,p)\supset D) = 1$ independently of p, provided n is large enough.

For all other digraphs, i.e. for those containing at least one directed cycle, we should expect two thresholds, one of '0–1'and the other of '1–0' type, since at the beginning, (when $p = 0$), and at the end of evolution (when $p = 1$) the tournament $T(n,p)$ is transitive.

To formulate our result we acquire the following notation. Let D be a digraph with a linear order imposed on its vertex set. Then we define $G(D)$ to be the undirected graph on the same vertex set as D, whose edges correspond to those arcs (i,j) of D for which $i < j$. For a graph $G = (V(G),E(G))$ let $d(G) = |E(G)|/|V(G)|$ and $m(G) = \max_{H\subseteq G} d(H)$.

Now, let D be an unlabelled digraph on k vertices. Among all copies of D on linearly ordered vertex set $[k]$, let D_0 be one which minimizes $m(G(D_0))$ and set $m(D) = m(G(D_0))$. For instance, $m(D) = \frac{1}{2}$ for any directed cycle D. Observe, that $m(D) = 0$ if and only if D is acyclic.

Theorem 1. *Let D be an antisymmetric digraph with at least one cycle. Then*

$$\lim_{n\to\infty} \mathrm{Prob}(T(n,p)\supset D) = \begin{cases} 0 & \text{if } np^{m(D)} \to 0, \\ 1 & \text{if } n(\min(p,1-p))^{m(D)} \to \infty, \\ 0 & \text{if } n(1-p)^{m(D)} \to 0. \end{cases}$$

Proof. We may restrict ourselves to the case $p \leqslant \frac{1}{2}$; the case $p \geqslant \frac{1}{2}$ follows by symmetry.

Note that $G(T(n,p)) = G(n,p)$, so the existence of a copy of D in $T(n,p)$ implies the existence of a copy of a graph H in $G(n,p)$ with $d(H) \geqslant m(D)$. But the expected number of such copies in $G(n,p)$ is $O(n^{|V(H)|}p^{|E(H)|}) = O((np^{d(H)})^{|V(H)|}) = o(1)$ whenever $np^{m(D)} \to 0$ as $n \to \infty$.

Assume now that $np^{m(D)} \to \infty$ and denote by X the number of those copies of D which are mapped into D_0 by some increasing bijection. The expectation of X is of the order $n^{|V(D)|}p^{|E(G(D_0))|}$ (due to the assumption that $p \leqslant \frac{1}{2}$ all powers of $1-p$ can be ignored). For the variance of X we have

$$\mathrm{Var}\, X = O\left(\sum_{D'} n^{2|V(D)|-|V(D')|} p^{2|E(G(D_0))|-|E(G(D'))|}\right),$$

where the sum is taken over all subdigraphs D' of D_0 with at least one arc. Hence

$$\mathrm{Prob}(X=0) \leqslant \frac{\operatorname{Var} X}{(\mathrm{E}X)^2} = \mathrm{O}\left(\sum_{D'} \frac{1}{n^{|V(D')|}p^{|E(G(D'))|}}\right) = \mathrm{o}(1),$$

since $|E(G(D'))|/|V(D')| \leqslant m(D)$. □

Remark. The second part of Theorem 1 can also be proved in the following way. Partition the vertex set of $T(n,p)$ into k consecutive blocks of size about n/k each and find an induced copy of $G(D_0)$ with the ith vertex in ith block. The existence of such a copy follows from the induced version of Corollary 1 from [4] (see also Remark 3 there).

Example. Let D be a directed cycle of length k. Then $G(D_0)$ is a graph on k vertices which contains at least one edge whose every connected component has at most two vertices. Thus $m(D) = \frac{1}{2}$ and, as long as $n^2 p \to 0$, there are almost surely no directed cycles in $T(n,p)$ (here and below *almost surely* means 'with probability tending to 1 as $n \to \infty$'). Indeed, for such p, tournament $T(n,p)$ almost surely contains no arcs (i,j) with $i < j$, so it is transitive. As soon as $n^2 p \to \infty$ (but $n^2(1-p) \to \infty$) there are almost surely many copies of D in $T(n,p)$. It can be proved (see Gruszka [3], Ruciński [7]) that then the number of copies of D is asymptotically normal. So far the situation is similar to the evolution of $G(n,p)$, where cycles of length k appear, independently of k, when $np \to \infty$. However, the threshold behaviours of these two models are quite different. If $np \to c > 0$, the number of cycles of length k in $G(n,p)$ tends to a random variable with Poisson distribution, whereas for $n^2 p \to c > 0$ the expectation of directed cycles of length k tends to infinity and their asymptotic distribution is more complicated.

The above example does not indicate that there is no room for the Poisson phase in the evolution of $T(n,p)$. It can be easily shown by the method of moments that, for $np^{m(D)} \to c > 0$, the distribution of the number of copies of D is asymptotically Poisson if and only if for every copy D_0 of D on a linearly ordered vertex set, for which $m(G(D_0)) = m(D)$, the graph $G(D_0)$ is strictly balanced. (An undirected graph G is strictly balanced if for every proper subgraph H of G we have $d(H) < d(G)$.) Let us call a digraph D which satisfies the above condition *strictly balanced*. Thus, the question arises whether such strictly balanced digraphs exist at all. The answer for this problem is positive and a strictly balanced digraph with 12 vertices was found by Andrzej Kurek who also proved that no digraph with less than 12 vertices is strictly balanced (personal communication).

3. Large structures

In this section we determine, for a wide range of $p = p(n)$, the size of the largest strong component and the length of the longest cycle in $T(n,p)$.

Call a tournament *irreducible* if it is not possible to partition its vertices into two non-empty sets A and B in such a way that no arc goes from A to B. It is well known (see [5]) that, for tournaments, the properties of being strongly connected, irreducible, Hamiltonian and pancyclic coincide – we shall use their equivalence a couple of times.

For simplicity we restrict our consideration to the case when $p \leqslant \frac{1}{2}$ – using the obvious duality between $T(n,p)$ and $T(n,1-p)$ one can easily extend our result to larger values of p.

Let

$$\alpha = \min\{i: (i,j) \in T(n,p) \text{ for some } j > i\}$$

and

$$\beta = \max\{j: (i,j) \in T(n,p) \text{ for some } i < j\},$$

while in an (unlikely) case when $(i,j) \notin T(n,p)$ for all $1 \leqslant i < j \leqslant n$ we put $\alpha = \beta = 1$. (We use Greek letters here to emphasize the fact that α and β are random variables.)

Clearly, each $i \in [n]$ such that either $i < \alpha$ or $i > \beta$ constitutes a one-element strongly connected component of its own. The main result of this section states that all vertices between α and β belong almost surely to one giant component of $T(n,p)$, provided only that $n^2 p \to \infty$.

Theorem 2. *If $n^2 p \to \infty$ and $p \leqslant 1/2$ then almost surely the set $\{\alpha, \alpha+1, \ldots, \beta\}$ is the vertex set of a strong component of $T(n,p)$.*

Proof. Let us consider first the case when $np \leqslant 3 \log n$. Our initial goal is to show that almost surely $T(n,p)$ contains a directed path from α to β.

Let $a = a(n) \leqslant b = b(n)$ be sequences of natural numbers such that $anp \to \infty$, $b^2 p \to \infty$ but $abp \to 0$. Note that these conditions imply that $1 \leqslant a = o(b)$ and $b = o(n)$. For instance, $a = 1 + \lfloor n^{-0.98} p^{-0.99} \rfloor$ and $b = \lfloor n^{0.02} p^{-0.49} \rfloor$ would do.

Let us split the set $[n]$ into five consecutive segments: $A_1 = \{1, \ldots, a\}$, $A_2 = \{a+1, \ldots, a+b\}$, $A_3 = \{a+b+1, \ldots, n-a-b\}$, $A_4 = \{n-a-b+1, \ldots, n-a\}$ and $A_5 = \{n-a+1, \ldots, n\}$. Thus, $|A_1| = |A_5| = a$, $|A_2| = |A_4| = b$ and $|A_3| = n-2a-2b = n - o(n)$. Note that the following facts hold almost surely:

(i) $\alpha \in A_1$ and $\beta \in A_5$.

(ii) $\forall i \in A_1: \forall j \in [n] \setminus A_3, j > i: (i,j) \notin T(n,p)$, and, similarly,

$$\forall i \in [n] \setminus A_3: \forall j \in A_5, j > i: (i,j) \notin T(n,p).$$

In particular, $\exists i_0, j_0 \in A_3: (\alpha, i_0), (j_0, \beta) \in T(n,p)$.

(iii) $\exists k_0 \in A_2: \exists l_0 \in A_4: (k_0, l_0) \in T(n,p)$.

Indeed, (i) and (iii) are true since the number of arcs from A_1 to $[n] \setminus A_1$, from $[n] \setminus A_5$ to A_5 and from A_2 to A_4 are binomially distributed with expectations of orders $anp \to \infty$, $anp \to \infty$ and $b^2 p \to \infty$, respectively. To see (ii) note that the expected number of arcs (i,j), with $i < j$, going either from A_1 to $A_1 \cup A_2 \cup A_4 \cup A_5$ or from $A_1 \cup A_2 \cup A_4 \cup A_5$ to A_5 is of order $abp \to 0$.

Now we use the standard method of generating a random graph in two rounds, fixing the outcome of the first round (we omit here a formal conditioning argument). In our case we first generate all arcs but those between A_2 and A_3 and between A_3 and A_4. Then, we find α, β, i_0, j_0, k_0 and l_0, and from now on we treat them as fixed vertices. Hence, the probability that in the second round both arcs (i_0, k_0) and (l_0, j_0) belong to $T(n,p)$ is $(1-p)^2 = 1 - o(1)$. Thus, almost surely $T(n,p)$ contains a directed path of length 5 (in our notation $\alpha i_0 k_0 l_0 j_0 \beta$) from α to β.

To prove the strong connectivity we need to find, for each $\alpha \leqslant j \leqslant \beta$, a path from β to j and a path from j to α. If $j = \alpha+1, \alpha+2, \beta-2, \beta-1$, the existence of such a path of length 1 follows from (i) and (ii). We claim that almost surely $T(n,p)$ contains, for each j, $\alpha+3 \leqslant j \leqslant \beta-3$, at least one of the paths $j\alpha$, $j(\alpha+1)\alpha$, $j(\alpha+2)\alpha$ and at least one of the paths βj, $\beta(\beta-1)j$ or $\beta(\beta-2)j$. Indeed, it follows from the fact that the expected number of quadruples of the form k, $k+1$, $k+2$, j which induce in $T(n,p)$ a subtournament with more than two arcs going 'to the right' is bounded from above by

$$n^2 2^6 p^3 = O(\log^3 n/n) = o(1).$$

Finally, assume that $3\log n \leqslant np \leqslant n/2$. Since for $np \to \infty$ almost surely $\alpha = 1$ and $\beta = n$ we must show that almost surely $T(n,p)$ is strongly connected, or, equivalently, that it is irreducible. Let X be the number of bipartitions of $[n]$ which violate irreducibility of $T(n,p)$. Then, the expectation of X is bounded from above by

$$\mathrm{E}X \leqslant \sum_{k=1}^{\lfloor n/2 \rfloor} \binom{n}{k} (\max\{1-p, p\})^{k(n-k)} \leqslant \sum_{k=1}^{\lfloor n/2 \rfloor} (n(1-p)^{n/2})^k = o(1)\,.$$

This completes the proof of Theorem 2. □

Thus we have shown that if $n^2 p \to \infty$ and $p \leqslant \frac{1}{2}$ then almost surely there are $\zeta = \alpha - 1 + n - \beta$ vertices $T(n,p)$ outside the giant component and the longest cycle is almost surely of the length $n - \zeta$ (note that the assumption $n^2 p \to \infty$ is necessary to assure that $T(n,p)$ is almost surely non-transitive.) It is not difficult to find the limit distribution of ζ. When $np \to 0$ we have

$$\lim_{n\to\infty} \mathrm{Prob}((\alpha-1)np \leqslant x, (n-\beta)np \leqslant y) = (1 - e^{-x})(1 - e^{-y})$$

so $np\zeta$ converges in distribution to the sum of independent, identically distributed exponential random variables. Consequently,

$$\lim_{n\to\infty} \mathrm{Prob}\left(\zeta \leqslant \frac{x}{np}\right) = 1 - (1+x)e^{-x} \quad \text{for } 0 < x < \infty\,.$$

In the threshold case $np \to c > 0$, for fixed i and j there are no arcs from $\{1, 2, \ldots, i\}$ to $\{n-j, n-j+1, \ldots, n\}$ so events '$\alpha = i$' and '$n - \beta = j$' are asymptotically independent. Thus, for $k = 0, 1, \ldots,$

$$\begin{aligned}\lim_{n\to\infty} \mathrm{Prob}(\zeta = k) &= \lim_{n\to\infty} \sum_{l=0}^{k} \mathrm{Prob}(\alpha - 1 = l, n - \beta = k - l)\\ &= \sum_{l=0}^{k} \lim_{n\to\infty} \mathrm{Prob}(\alpha - 1 = l) \lim_{n\to\infty} \mathrm{Prob}(n - \beta = k - l)\\ &= \sum_{l=0}^{k} (\mathrm{e}^{-cl} - \mathrm{e}^{-c(l+1)})(\mathrm{e}^{-c(k-l)} - \mathrm{e}^{-c(k-l+1)})\\ &= (k+1)(1 - \mathrm{e}^{-c})^2 \mathrm{e}^{-ck}.\end{aligned}$$

In particular, if $p \leqslant \frac{1}{2}$, then

$$\lim_{n\to\infty} \mathrm{Prob}(T(n,p) \text{ is pancyclic}) = \lim_{n\to\infty} \mathrm{Prob}(T(n,p) \text{ is Hamiltonian})$$

$$= \lim_{n\to\infty} \mathrm{Prob}(T(n,p) \text{ is strongly connected}) = \begin{cases} 0 & \text{if} \quad np \to 0, \\ (1 - \mathrm{e}^{-c})^2 & \text{if} \quad np \to c, \\ 1 & \text{if} \quad np \to \infty. \end{cases}$$

For $p = \frac{1}{2}$, the above result was already shown in [6], a proof of the case when $np \to \infty$ was first presented in [3].

References

[1] O. Frank, Stochastic competition graphs, Rev. Int. Statist. Inst. 36 (1971) 319–326.

[2] R. Graham, B. Rothschild and J. Spencer, Ramsey Theory (Wiley, New York, 1990)

[3] J. Gruszka, Cycles in tournaments and random tournaments. Ph.D. Dissertation, Adam Mickiewicz University, Poznań, 1990 (in Polish).

[4] M. Karoński and A. Ruciński, Small subgraphs of k-partite random graphs, Combin. Math.; Proc. 3rd Internat. Conf. (The New York Academy of Sciences, New York, 1989) 230–240.

[5] J.W. Moon, Topics on Tournaments (Holt, Rinehart and Winston, New York, 1968).

[6] J.W. Moon and L. Moser, Almost all tournaments are irreducible, Can. Math. Bull. 5 (1962) 61–65.

[7] A. Ruciński, Proving normality in combinatorics, in: A. Frieze and T. Luczak, eds., Random Graphs, Vol. 2 (Wiley, New York 1992) 215–232.